

---

---

# **Algebraic multilevel preconditioner with projectors**

**Konstantin Lipnikov**

**Yuri Kuznetsov**

Los Alamos National Laboratory  
MS B284, Los Alamos, NM 87545  
[lipnikov@t7.lanl.gov](mailto:lipnikov@t7.lanl.gov)

Department of Mathematics, University of Houston  
Houston, TX 77204, [kuz@math.uh.edu](mailto:kuz@math.uh.edu)

Acknowledgments: research has been supported by LACSI; special thanks to D.Moulton

# Contents

---

---

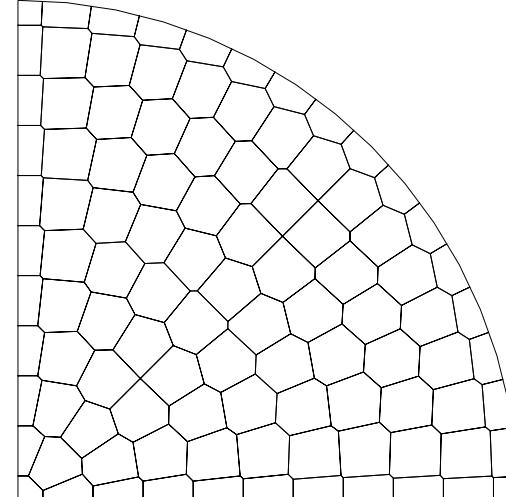
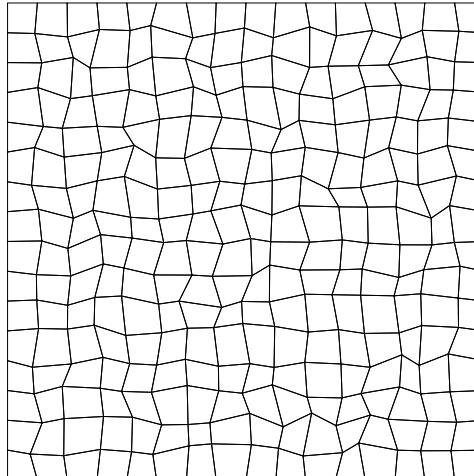
- Motivation
- Model problem
- Two-level preconditioner
- Multi-level preconditioner
- Numerical comparison of three MG preconditioners
- Merging different MG ideas
- Numerical experiments
- Conclusion

# Motivation

---

We want to attack the following problems:

- problems resulting in stiff SPD systems (in particular, SPD systems for edge-based intensities in mixed FE methods);
- problems with materials having fine heterogeneous structure (e.g., different diffusion coefficients in every mesh cell);
- problems on structured and unstructured meshes having elements of mixed types (e.g., Voronoï meshes).



# Model problem (1/3)

---

For simplicity, we consider the elliptic equation

$$-\operatorname{div}(K \nabla u) = b \quad \text{in } \Omega$$

with Neumann b.c. on  $\partial\Omega$  and the compatibility condition

$$\int_{\Omega} b \, dx = 0.$$

We rewrite the problem as a system of first order equations:

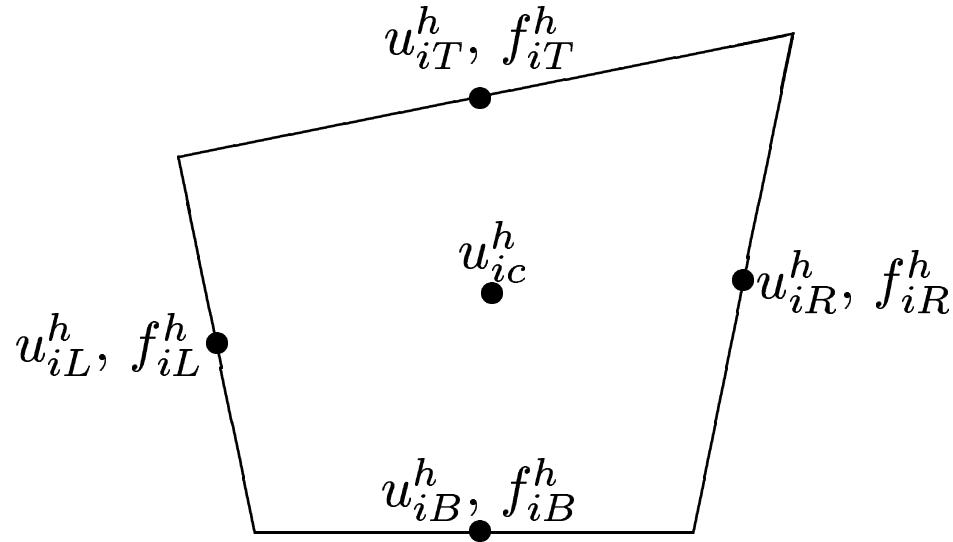
$$\operatorname{div} \mathbf{f} = b, \quad \mathbf{f} = -K \nabla u.$$

We consider topologically rectangular grids.

# Model problem (2/3)

---

We discretize the problem with the Raviart-Thomas finite elements:



In each mesh cell we get the discrete equations:

$$(\text{DIV } \mathbf{f}^h)_i = b_{ic}, \quad \mathbf{f}_i^h = (\text{GRAD } u^h)_i,$$

where  $i = 1, \dots, m$ .

# Model problem (3/3)

The global discretization is achieved by imposing the continuity of fluxes

$$f_{iR}^h = -f_{jL}^h$$

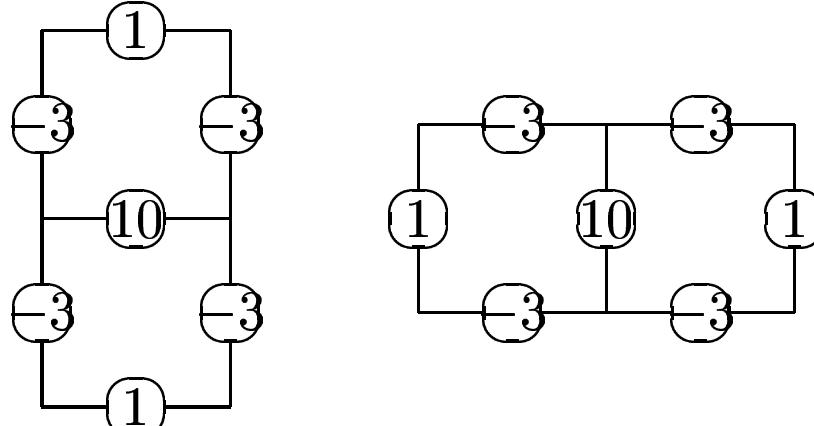
and interface (edge-based) intensities

$$u_{iR}^h = u_{jL}^h.$$

Eliminating fluxes and cell-centered intensities, we get an algebraic problem

$$\textcolor{blue}{A}u = b$$

for edge-based intensities. For a square grid the stencil of  $\textcolor{blue}{A}$  is as follows:



# Two-level preconditioner (1/7)

---

Matrix  $\mathcal{A}$  can be written in an assembling form

$$\mathcal{A} = \sum_{i=1}^m \mathcal{N}_i A_i \mathcal{N}_i^T$$

where  $A_i$  is an elemental matrix and  $\mathcal{N}_i$  is an assembling matrix.

The preconditioner  $\mathcal{B}$  is given by

$$\mathcal{B} = \sum_{i=1}^m \mathcal{N}_i B_i \mathcal{N}_i^T$$

where

$$B_i \sim A_i.$$

# Two-level preconditioner (2/7)

---

Consider an eigenvalue problem

$$A_i w = \lambda M_i w, \quad M_i = M_i^T > 0.$$

Then,

$$A_i = M_i W_i \Lambda_i W_i^T M_i$$

where

$$\Lambda_i = \text{diag}\{\lambda_1^i, \dots, \lambda_s^i\}, \quad \lambda_1^i = 0,$$

and  $W_i$  is a matrix of eigenvectors. Define

$$B_i = M_i W_i \tilde{\Lambda}_i W_i^T M_i$$

where

$$\tilde{\Lambda}_i = \text{diag}\{0, 1, \dots, 1\}.$$

# Two-level preconditioner (3/7)

---

- Let  $M_i = \text{diag}(A_i)$  (another choice is a lumped mass matrix);
- Then,  $B_i$  is close to  $A_i$  in a sense that,

$$\lambda_2^i(B_i x, x) \leq (A_i x, x) \leq \lambda_4^i(B_i x, x) \quad \forall x \in \mathbb{R}^4,$$

and  $\lambda_4^i/\lambda_2^i$  depends only on the element shape and tensor anisotropy;

- Define  $P_i = w_1^i [w_1^i]^T$ ,

$$P_i = \alpha_i \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix}.$$

Then,

$$B_i = M_i - M_i P_i M_i.$$

# Two-level preconditioner (4/7)

---

**Theorem.**

$$c_1(\mathbf{B}x, x) \leq (\mathbf{A}x, x) \leq c_2(\mathbf{B}x, x) \quad \forall x \in \Re^n$$

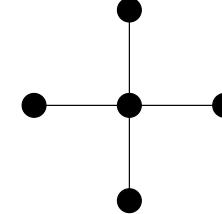
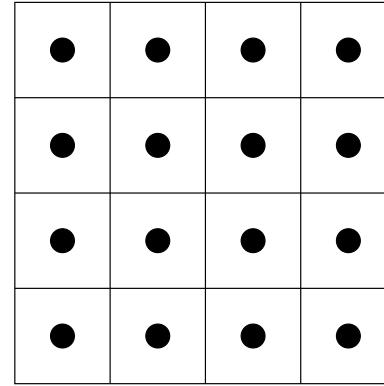
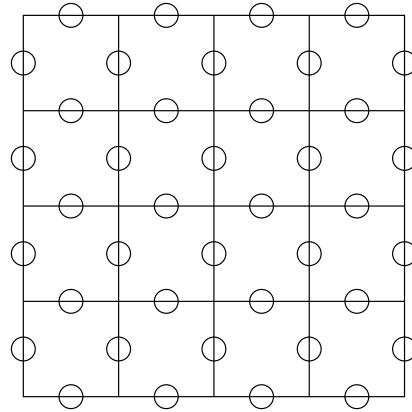
where  $c_1$  and  $c_2$  depend only on the shape of mesh cells and tensor anisotropy.

$$\begin{aligned}\mathbf{B} = \sum_{i=1}^m \mathcal{N}_i B_i \mathcal{N}_i^T &= \sum_{i=1}^m \mathcal{N}_i (M_i - M_i P_i M_i) \mathcal{N}_i^T \\ &= \mathbf{M} - \sum_{i=1}^m \mathcal{N}_i M_i P_i M_i \mathcal{N}_i^T \equiv \mathbf{M} - \mathbf{C}.\end{aligned}$$

- $\text{rank } \mathbf{C} = m < n = \text{rank } \mathbf{B}$       (<#cells < # edges>);
- $\mathbf{B}$  is a  $M$ -matrix;
- $P_i M_i$  is  $M_i$ -orthogonal projector onto kernels of both  $A_i$  and  $B_i$ .

# Two-level preconditioner (5/7)

---



Consider the system

$$\mathbf{B}v = (\mathbf{M} - \mathbf{C})v = g.$$

Define

$$\hat{v}_i = [w_1^i]^T M_i \mathcal{N}_i^T v \quad \text{and} \quad \hat{g}_i = [w_1^i]^T M_i \mathcal{N}_i^T \mathbf{M}^{-1} g.$$

If  $M_i$  is a mass matrix, then

$$\hat{v}_i = [w_1^i]^T M_i \mathcal{N}_i^T v = \frac{1}{\|e_i\|} \int_{e_i} v^h \, dx.$$

# Two-level preconditioner (5/7)

---

Let us multiply the equation

$$\textcolor{blue}{B}v = \textcolor{blue}{M}v - \textcolor{blue}{C}v = g$$

by

$$[w_1^i]^T M_i \mathcal{N}_i^T \textcolor{blue}{M}^{-1}, \quad i = 1, \dots, m.$$

We get the following SLAEs:

$$(I - \hat{Q})\hat{v} = \hat{g}, \quad \hat{Q} \in \Re^{m \times m},$$

where

$$\hat{Q} = (\hat{q}_{ij}) \quad \hat{q}_{ij} = [w_1^i]^T M_i \mathcal{N}_i^T \textcolor{blue}{M}^{-1} \mathcal{N}_j M_j w_1^j.$$

- $I - \hat{Q}$  is a symmetric positive semi-definite matrix;
- $I - \hat{Q}$  is  $M$ -matrix.

# Two-level preconditioner (7/7)

---

Let  $Z = (I - \hat{Q})^+ = (z_{ij})$ . Then

$$\mathbf{B}^+ = \mathbf{M}^{-1} + \mathbf{M}^{-1} \left[ \sum_{i,j=1}^m z_{ij} \mathcal{N}_i^T M_i w_i [w_j]^T M_j \mathcal{N}_j \right] \mathbf{M}^{-1}.$$

**Theorem.** If

$$R \sim Z, \quad \ker R = \ker Z,$$

and

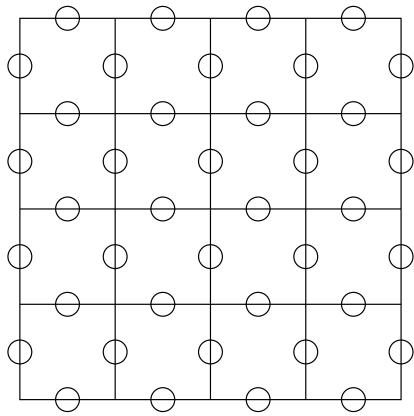
$$\mathbf{B}_R^+ = \mathbf{M}^{-1} + \mathbf{M}^{-1} \left[ \sum_{i,j=1}^m r_{ij} \mathcal{N}_i^T M_i w_i [w_j]^T M_j \mathcal{N}_j \right] \mathbf{M}^{-1}$$

then,

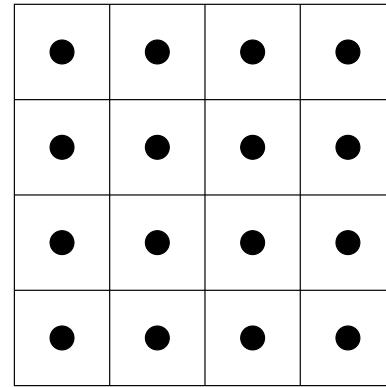
$$\mathbf{B}_R \sim \mathbf{B}.$$

In practice, it means that instead of  $(I - \hat{Q}) \hat{v} = \hat{g}$  we compute  $\hat{v} = R \hat{g}$ .

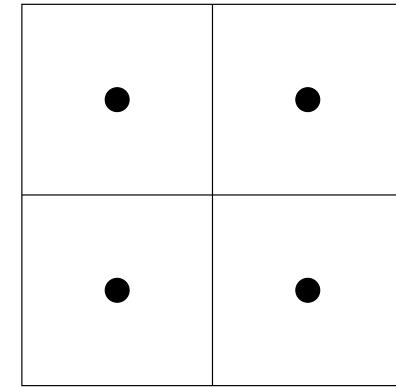
# Multi-level preconditioner (1/3)



$A$



$B = A_1 = M - C_1$



$A_2 = M - C_2$

Recall, that

$$A_1 = \sum_{i=1}^m \mathcal{N}_{1,i} (M_{1,i} - M_{1,i} P_{1,i} M_{1,i}) \mathcal{N}_{1,i}^T.$$

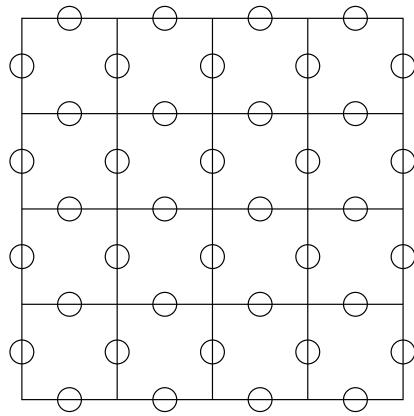
Define

$$A_2 = \sum_{i=1}^{m/4} \mathcal{N}_{2,i} (M_{2,i} - M_{2,i} P_{2,i} M_{2,i}) \mathcal{N}_{2,i}^T$$

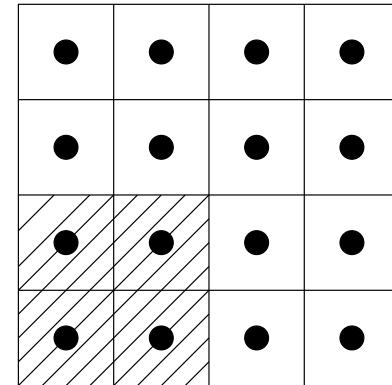
where

$$M_{2,i} = \sum_j \hat{\mathcal{N}}_{1,j} M_{1,j} \hat{\mathcal{N}}_{1,j}^T.$$

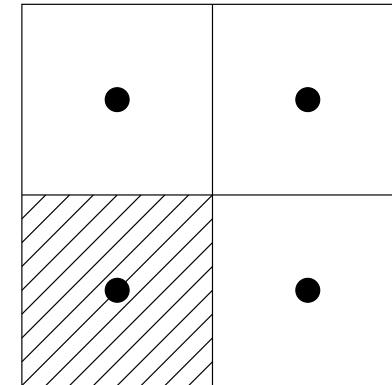
# Multi-level preconditioner (1/3)



$A$



$B = A_1 = M - C_1$



$A_2 = M - C_2$

Recall, that

$$A_1 = \sum_{i=1}^m \mathcal{N}_{1,i} (M_{1,i} - M_{1,i} P_{1,i} M_{1,i}) \mathcal{N}_{1,i}^T.$$

Define

$$A_2 = \sum_{i=1}^{m/4} \mathcal{N}_{2,i} (M_{2,i} - M_{2,i} P_{2,i} M_{2,i}) \mathcal{N}_{2,i}^T$$

where

$$M_{2,i} = \sum_j \hat{\mathcal{N}}_{1,j} M_{1,j} \hat{\mathcal{N}}_{1,j}^T.$$

# Multi-level preconditioner (2/3)

---

Repeating the coarsening algorithm, we get a sequence of full rank matrices  $\mathbf{A}, \mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_L$  such that

$$\mathbf{A}_l = \mathbf{M} - \mathbf{C}_l, \quad l = 1, \dots, L.$$

Define space  $V_l = \text{im} \mathbf{C}_l$ . Then

- $\dim V_l = \text{rank } \mathbf{C}_l = m_l \approx m_{l-1}/4$ ;
- $V_L \subset V_{L-1} \subset \dots \subset V_1 \subset \mathbb{R}^n$ ;

**Lemma.** Let  $\mathcal{T}_{l,h}$  be a sequence of square nested meshes. Then

$$\frac{c_1}{2^l} (\mathbf{A}_l x, x) \leq (\mathbf{A}_{l-1} x, x) \leq c_2 (\mathbf{A}_l x, x) \quad \forall x \in \mathbb{R}^n$$

where  $c_1$  and  $c_2$  are independent of  $l$  and the discretization parameter  $h$ .

# Multi-level preconditioner (3/3)

---

In order to evaluate  $\tilde{r} = \textcolor{red}{H}_l^{-1}r$ ,  $r \in V_{l-1}$ , we proceed as follows:

- take a special initial guess:

$$x^0 = \textcolor{blue}{M}^{-1}r;$$

- enter a subspace:

$$\xi^0 = \textcolor{blue}{A}_l x^0 - r = -\textcolor{blue}{C}_l \textcolor{blue}{M}^{-1}r, \quad y^0 = -\xi^0, \quad \xi^0 \in V_l;$$

- iterate in the subspace :

$$\begin{aligned}\xi^i &= \xi^{i-1} - \gamma_i \textcolor{blue}{A}_l \textcolor{red}{H}_{l+1}^{-1} \xi^{i-1} \\ y^i &= y^{i-1} - \gamma_i \textcolor{blue}{C}_l \textcolor{red}{H}_{l+1}^{-1} \xi^{i-1}, \quad i = 1, \dots, \beta;\end{aligned}$$

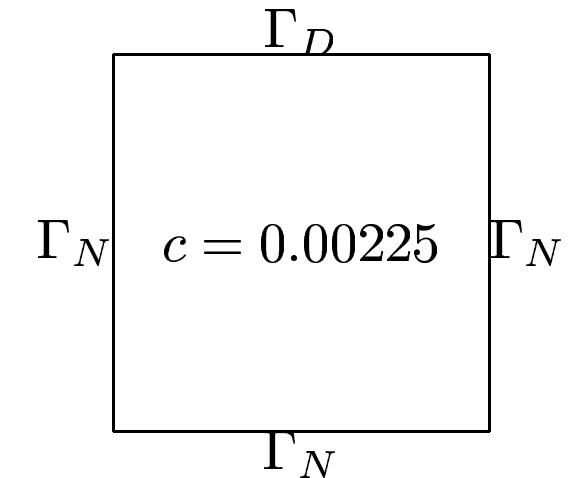
- leave the subspace:

$$\tilde{r} = \textcolor{blue}{M}^{-1}(y^\beta + r).$$

# Numerical experiments (1/4)

---

$$\begin{aligned}-\operatorname{div}(K \nabla p) + c p &= 1 && \text{in } \Omega \\ p &= 0 && \text{on } \Gamma_D \\ (K \nabla p) \mathbf{n} &= 0 && \text{on } \Gamma_N\end{aligned}$$



Components used in numerical experiments:

- W(0,0) cycle of Multilevel preconditioner with projectors;
- W(1,1) cycle of Black Box MG (J.Dendy, D.Moulton, release 1.5.2, 2001);
- V(1,1) cycle of AMG (J.Ruge, K.Stüben, R.Hempel, release 1.5, 1990).

# Numerical experiments (2/4)

---

The test problems are:

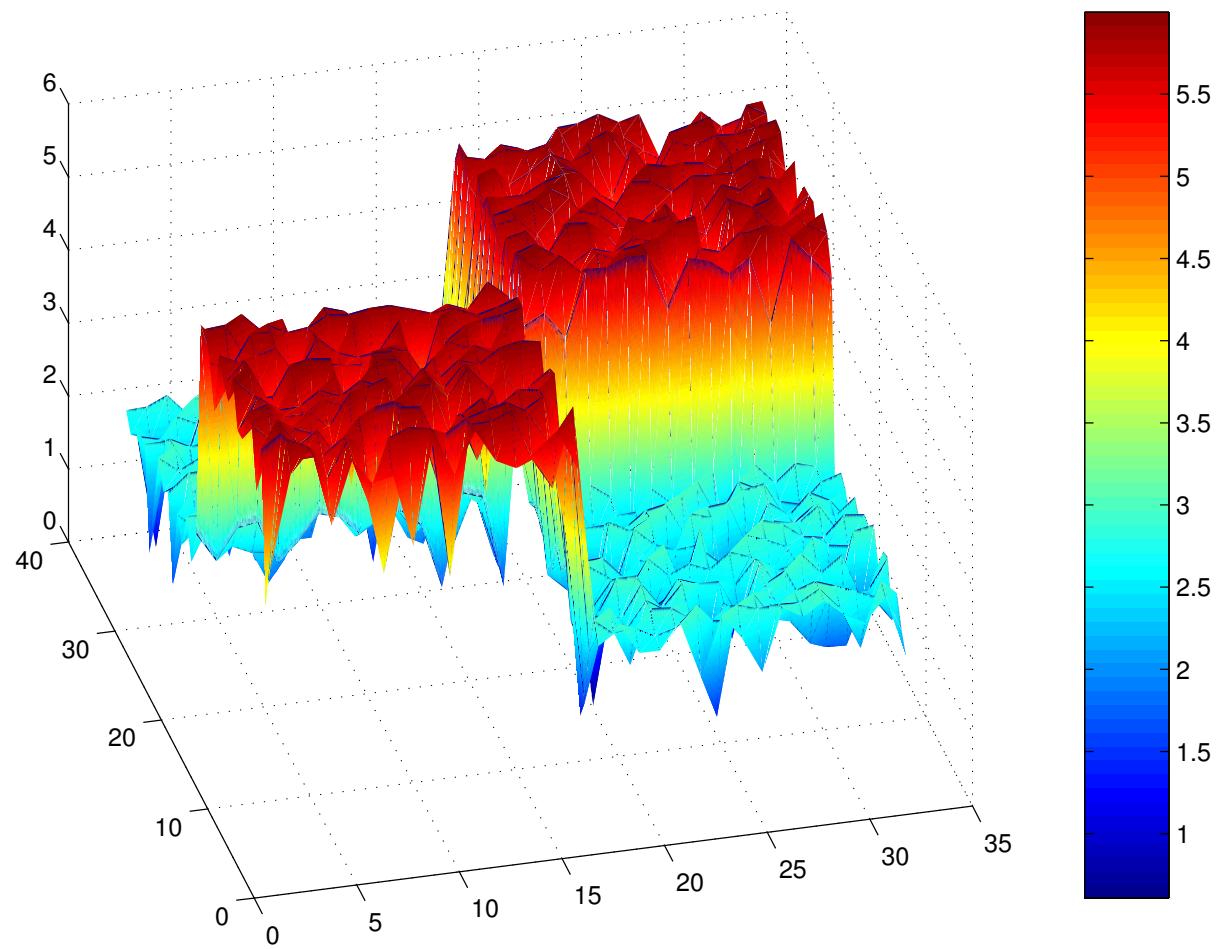
1.  $K = K_1 = 1$ ;
2.  $K = K_2(x)$  where

$$K_2 = \begin{cases} 10^3, & x \in (0; 0.5)^2 \cup (0.5; 1)^2 \\ 1, & \text{otherwise;} \end{cases}$$

3.  $K = K_3 = K_2 \psi(i)$ ,  $1 \leq \psi \leq 10^3$ ,  $i = 1, \dots, m$ .

# Numerical experiments (3/4)

---

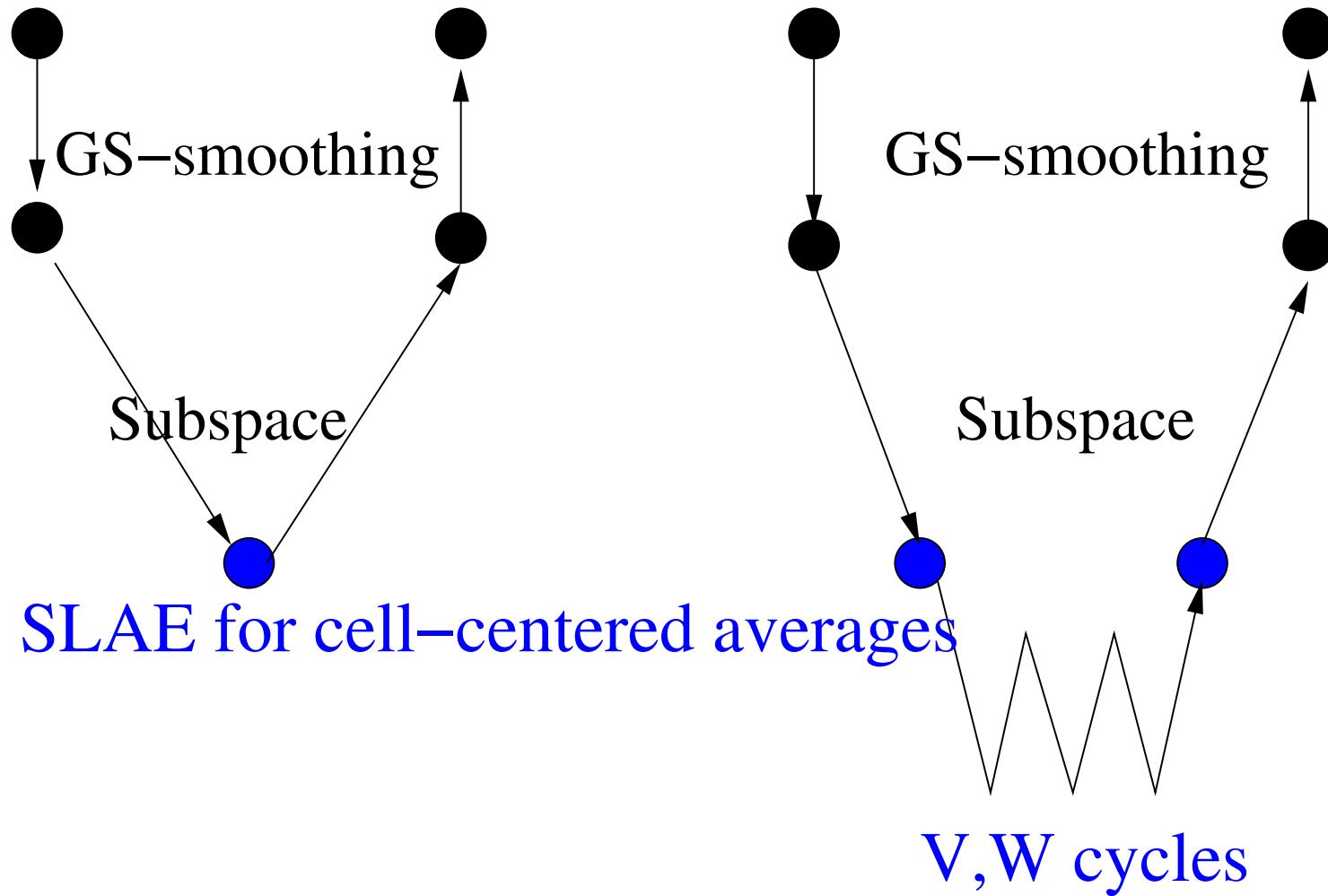


Example of  $K_3(i, j)$ .

# Numerical experiments (4/4)

---

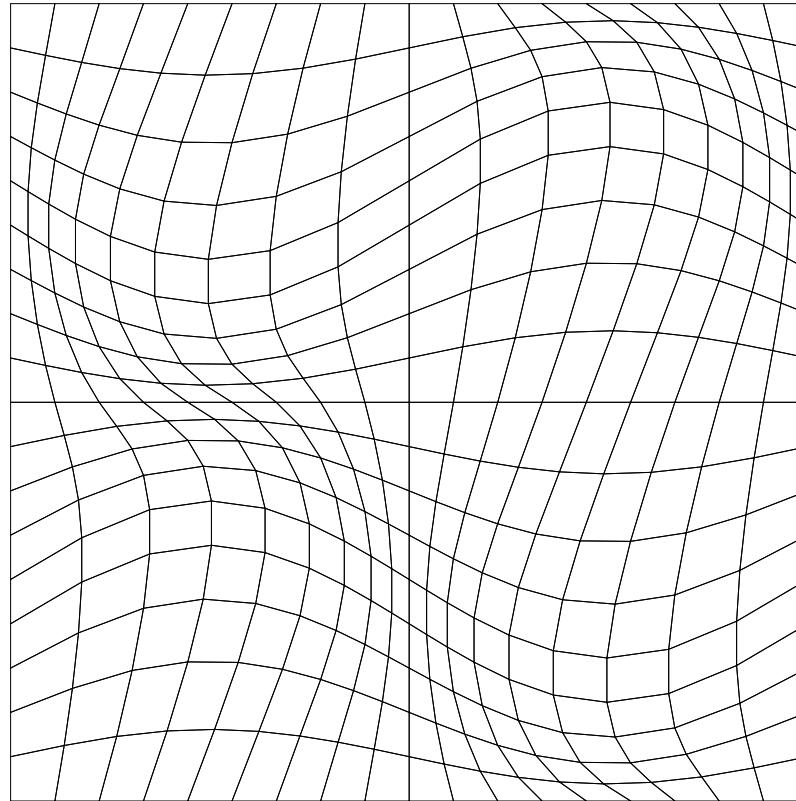
SLAE for edge-based intensities



# Smooth grids (1/4)

---

Grids are generated via a smooth mapping of reference rectangular grids:



# Smooth grids (2/4)

---

- We consider both mimetic FD and mixed-hybrid FE discretizations. In the case of square grids, the elemental matrices are given by:

$$A_i^{FD} = \begin{bmatrix} 3 & -1 & -1 & 0 \\ -1 & 3 & 0 & -1 \\ -1 & 0 & 3 & -1 \\ 0 & -1 & -1 & 3 \end{bmatrix}, \quad A_i^{FE} = \begin{bmatrix} 5 & 1 & -3 & -3 \\ 1 & 5 & -3 & -3 \\ -3 & -3 & 5 & 1 \\ -3 & -3 & 1 & 5 \end{bmatrix}.$$

- We compare the number of PCG iterations for relative tolerance  $\varepsilon = 10^{-6}$  and the **total** solution time.

# Smooth grids: Mimetic FD

		Jacobi		Dendy		AMG		Kuznetsov	
	1/h	# itr	time	# itr	time	# itr	time	# itr	time
$K_1$	12	78	0.01	—	—	8	0.01	11	0.02
	24	167	0.14	14	0.08	10	0.08	14	0.09
	48	359	1.54	17	0.39	10	0.44	17	0.53
	96	749	15.4	20	2.03	12	2.21	20	2.94
	192	1547	130.	22	9.64	13	9.57	23	15.0
$K_2$	12	93	0.02	—	—	8	0.01	11	0.02
	24	195	0.20	18	0.09	11	0.08	15	0.09
	48	411	1.79	22	0.49	12	0.49	18	0.57
	96	862	16.8	28	2.73	12	2.17	22	3.18
	192	1776	149.	33	13.8	16	10.9	26	16.7
$K_3$	12	114	0.02	—	—	8	0.02	11	0.03
	24	236	0.14	27	0.12	11	0.08	14	0.11
	48	528	2.16	49	0.97	14	0.49	21	0.71
	96	1059	20.5	89	7.93	19	2.95	25	3.98
	192	2191	185.	181	70.3	22	14.4	28	20.5

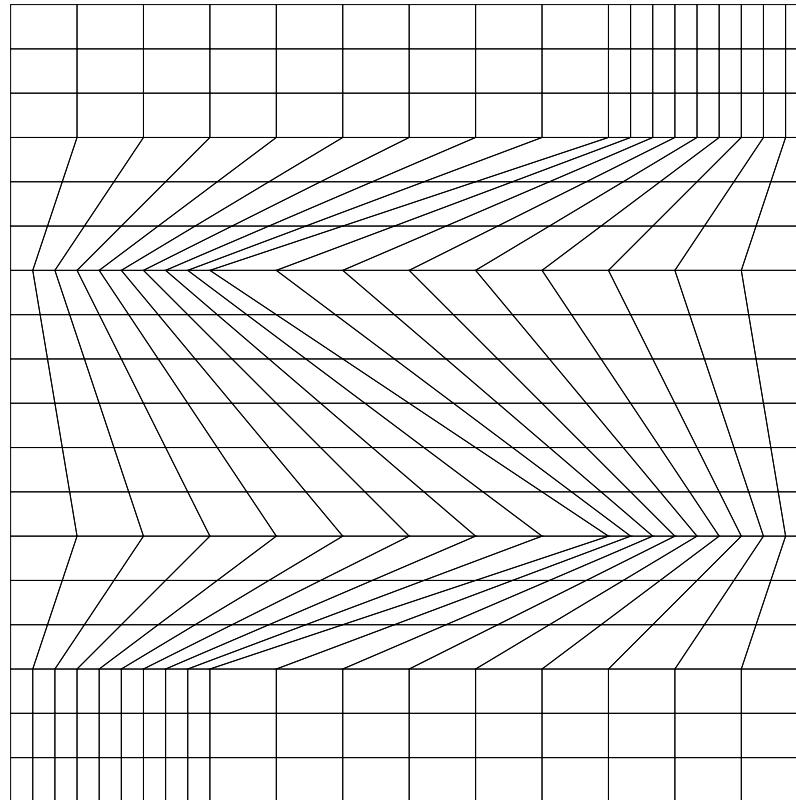
# Smooth grids: Mixed-hybrid FE

		Jacobi		Dendy		AMG		Kuznetsov	
	1/h	# itr	time	# itr	time	# itr	time	# itr	time
$K_1$	12	94	0.02	–	–	11	0.03	143	0.02
	24	194	0.24	16	0.10	12	0.12	17	0.12
	48	403	1.95	20	0.50	17	0.69	20	0.70
	96	828	17.9	23	2.55	25	4.04	22	3.54
	192	1706	158.	26	12.9	37	22.4	26	19.3
$K_2$	12	113	0.03	–	–	11	0.01	14	0.03
	24	231	0.30	21	0.12	13	0.09	17	0.12
	48	472	2.27	26	0.63	17	0.59	21	0.73
	96	960	20.8	30	3.21	26	3.70	24	3.86
	192	1964	180.	37	16.8	42	22.9	29	20.3
$K_3$	12	128	0.03	–	–	8	0.02	14	0.03
	24	267	0.35	33	0.17	12	0.16	16	0.12
	48	577	2.75	58	1.27	17	0.70	21	0.72
	96	1146	25.0	109	10.7	23	4.14	25	3.99
	192	2315	213.	217	90.5	25	23.1	29	20.3

# Kershaw grids (1/3)

---

Let us consider grids which are close to grids we can expect in real applications:



# Kershaw grids: Mimetic FD

		Jacobi		Dendy		AMG		Kuznetsov	
	1/h	# itr	time	# itr	time	# itr	time	# itr	time
$K_1$	12	101	0.03	–	–	12	0.02	17	0.04
	24	212	0.25	24	0.12	16	0.10	26	0.17
	48	438	2.11	30	0.70	18	0.65	32	1.03
	96	909	19.8	37	3.87	27	4.17	40	6.00
	192	1873	172.	45	20.1	26	16.4	49	32.5
$K_2$	12	138	0.03	–	–	13	0.02	23	0.04
	24	275	0.32	37	0.17	20	0.13	35	0.22
	48	536	2.60	45	0.99	23	0.80	38	1.19
	96	1087	23.6	53	5.35	25	3.89	49	7.27
	192	2202	202.	61	25.7	28	17.8	58	38.1
$K_3$	12	156	0.04	–	–	18	0.02	23	0.04
	24	331	0.41	44	0.21	21	0.13	34	0.22
	48	679	3.27	65	1.40	30	0.96	44	1.39
	96	1354	29.2	111	10.8	40	5.64	60	8.65
	192	2700	248.	211	88.3	43	25.1	66	43.1

# Kershaw grids: Mixed-hybrid FE

		Jacobi		Dendy		AMG		Kuznetsov	
	1/h	# itr	time	# itr	time	# itr	time	# itr	time
$K_1$	12	103	0.03	–	–	16	0.02	17	0.03
	24	219	0.26	23	0.12	27	0.18	23	0.16
	48	463	2.23	30	0.70	49	1.51	29	0.94
	96	969	21.0	37	3.90	101	13.4	34	5.22
	192	2006	185.	45	20.1	209	110.	42	28.2
$K_2$	12	136	0.03	–	–	22	0.03	22	0.04
	24	276	0.36	36	0.18	33	0.21	30	0.19
	48	570	2.77	44	0.98	57	1.71	36	1.12
	96	1183	25.8	51	5.18	114	15.0	43	6.40
	192	2418	222.	62	27.2	232	123.	51	33.6
$K_3$	12	163	0.04	–	–	13	0.02	23	0.04
	24	332	0.44	43	0.21	27	0.16	30	0.20
	48	714	3.47	67	1.46	79	2.21	39	1.22
	96	1460	36.1	123	12.0	151	19.4	49	7.23
	192	2946	270.	244	102.	299	157.	57	37.6

# Merging different MGs (1/3)

---

Recall that the SLAE for edge-based intensities

$$\textcolor{blue}{B}v = g, \quad g \in \Re^n,$$

is reduced to the SLAE for cell-centered averages:

$$(I - \hat{Q})\hat{v} = \hat{g}, \quad \hat{g} \in \Re^m.$$

- $I - \hat{Q}$  is positive semi-definite; however, it usually lacks diagonal dominance;
- $\ker(I - \hat{Q})$  is represented by a vector  $\hat{e} \in \Re^m$  with strongly oscillating entries.

# Merging different MGs (2/3)

---

A possible remedy is to rescale matrix  $I - \hat{Q}$ . Let us define a diagonal matrix  $D$  s.t.

$$\hat{e} = D \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Then, equation

$$(I - \hat{Q})\hat{v} = \hat{g}$$

can be replaced by an equivalent equation

$$D(I - \hat{Q})D \tilde{v} = D \hat{g}, \quad \hat{v} = D \tilde{v}.$$

- matrix  $\tilde{Q} = D(I - \hat{Q})D$  has a weak diagonal dominance.

# Merging different MGs (3/3)

		Dendy		Dendy++		AMG		AMG++		Kuznetsov	
	1/h	# itr	time	# itr	time	# itr	time	# itr	time	# itr	time
$K_1$	12	–	–	–	–	16	0.02	–	–	17	0.03
	24	23	0.12	23	0.12	27	0.18	23	0.14	23	0.16
	48	30	0.70	27	0.64	49	1.51	27	0.71	29	0.94
	96	37	3.90	32	3.37	101	13.4	33	3.81	34	5.22
	192	45	20.1	35	15.9	209	110.	37	17.3	42	28.2
$K_2$	12	–	–	–	–	22	0.03	–	–	22	0.04
	24	36	0.18	28	0.14	33	0.21	28	0.16	30	0.19
	48	44	0.98	33	0.76	57	1.71	34	0.85	36	1.12
	96	51	5.18	39	4.06	114	15.0	39	4.74	43	6.40
	192	62	27.2	45	20.0	232	123.	46	21.1	51	33.6
$K_3$	12	–	–	–	–	13	0.02	–	–	23	0.04
	24	43	0.21	29	0.15	27	0.16	29	0.17	30	0.20
	48	67	1.46	36	0.82	79	2.21	36	0.93	39	1.22
	96	123	12.0	41	4.22	151	19.4	41	4.71	49	7.23
	192	244	102.	47	21.1	299	157.	48	22.8	57	37.6

# Conclusion

---

---

1. In comparison with other preconditioners, the multilevel preconditioner was more robust on highly distorted meshes. Even better performance may be achieved with more relaxation sweeps on coarser grids.
2. The best preconditioner was obtained when we combined two-grid ML preconditioner with BoxMG as a coarse grid preconditioner. The proper scaling of the coarse grid matrix is crucial.
3. The method is easily generalized to unstructured grids. In this case the AMG should be used as a coarse grid preconditioner.
4. The geometric nature of the ML preconditioner is very convenient for its parallel implementation.
5. A quality of the ML preconditioner depends on the tensor anisotropy.